CS 7545: Machine Learning Theory	**Fall 2021**

Lecture 5

Name: Tyler LaBonte **Professor:** Santosh S. Vempala

1 Independent Component Analysis

Suppose we have samples $\boldsymbol{x} = \mathbf{A}\boldsymbol{s}$ where $\mathbf{A} \in \mathbb{R}^{n \times d}$ is unknown and $\boldsymbol{s} \in \mathbb{R}^d$ has independent coordinates. We can think of the "cocktail party" problem where we have d signals and n observations, and we'd like to transform the mixed signals into their maximally independent components via a linear transformation.

We can assume $\mathbb{E}[\boldsymbol{s}] = 0$ and $\mathbb{E}[s_i^2] = 1$ since we can recenter and rescale $\mathbf{A}$. We'd like to recover the columns of $\mathbf{A}$ up to sign. Notice this is not necessarily uniquely identifiable: if $s_1, s_2 \sim \mathcal{N}(0, 1)$, then for any rotation $\mathbf{A}$ we have $\boldsymbol{x} = \mathbf{A}\boldsymbol{s} \sim \mathcal{N}(0, \mathbf{I}_2)$.

Fact 1.1: Unique Identifiability of ICA

If at most one component of $\mathbf{A}$ is Gaussian, then the columns of $\mathbf{A}$ are uniquely identifiable up to sign.

Let's try moments. We have

$$\mathbb{E}[\boldsymbol{x}] = \mathbb{E}[\mathbf{A}\boldsymbol{s}] = \mathbf{A}\mathbb{E}[\boldsymbol{s}] = 0 \quad \mathbb{E}[\boldsymbol{x}\boldsymbol{x}^\intercal] = \mathbf{A}\mathbb{E}[\boldsymbol{s}\boldsymbol{s}^\intercal]\mathbf{A}^\intercal = \mathbf{A}\mathbf{A}^\intercal. \tag{1}$$

This doesn't help us, and $\mathbb{E}[\otimes^3 \boldsymbol{x}]$ might be 0 (*e.g.,* if the s_i are symmetric). Let's try fourth moments:

$$\begin{aligned}
\mathbb{E}[\otimes^4 \boldsymbol{x}]_{i,j,k,l} &= \mathbb{E}[(\mathbf{A}\boldsymbol{s})_i](\mathbf{A}\boldsymbol{s})_j(\mathbf{A}\boldsymbol{s})_k(\mathbf{A}\boldsymbol{s})_l] && (2)\\
&= \mathbb{E}[\sum_{i'} A_{ii'}s_{i'} \sum_{j'} A_{jj'}s_{j'} \sum_{k'} A_{kk'}s_{k'} \sum_{l'} A_{ll'}s_{l'}] && (3)\\
&= \sum_{i',j',k',l'} A_{ii'}A_{jj'}A_{kk'}A_{ll'}\mathbb{E}[s_{i'}s_{j'}s_{k'}s_{l'}]. && (4)
\end{aligned}$$

This expectation is $\mathbb{E}[s_{i'}^4]$ if each index is distinct, $\mathbb{E}[s_i^2]\mathbb{E}[s_k^2]$ if there are two pairs, and zero otherwise. Let

$$M_{ijkl} = \mathbb{E}[x_i x_j]\mathbb{E}[x_k x_l] + \mathbb{E}[x_i x_k]\mathbb{E}[x_j x_l] + \mathbb{E}[x_i x_l]\mathbb{E}[x_j x_k], \tag{5}$$

then

$$= \sum_{i'} A_{ii'}A_{ji'}A_{ki'}A_{li'}\mathbb{E}[s_i^4 - 3] + M_{ijkl}. \tag{6}$$

Therefore,

$$\mathbb{E}[\otimes^4 \boldsymbol{x}] - \mathbf{M} = \sum_i (\mathbb{E}[s_i^4] - 3) \otimes^4 \mathbf{A}_i. \tag{7}$$

This is a tensor decomposition into orthogonal vectors, so we can solve it by tensor power iteration!